MAC-315 - Programação Linear
Primeiro semestre de 2008
Prof. Marcelo Queiroz
http://www.ime.usp.br/~mqz

Notas de Aula[1]

1 Introdução

1.1 Exemplos de problemas de programação linear

O problema da dieta

Considere n diferentes alimentos e m diferentes nutrientes, e suponha que você possue uma tabela com o conteúdo nutricional de uma unidade ou porção de cada alimento:

	alim. 1	$\cdots$	**alim. n**
nutr. 1	a_{11}	$\cdots$	a_{1n}
$\vdots$	$\vdots$		$\vdots$
nutr. 1	a_{m1}	$\cdots$	a_{mn}

Note que a j-ésima coluna da matriz representa o conteúdo nutricional do j-ésimo alimento. Seja b_i o requisito nutricional mínimo do nutriente i em uma dieta balanceada. Podemos interpretar um vetor $x \in \mathbb{R}^n_+$ como a especificação de uma dieta que utiliza x_j unidades/porções do alimento j para cada $j = 1, \ldots, n$. A dieta x será balanceada se satisfizer $\sum_{j=1}^{n} a_{ij}x_j \geq b_i$ para cada nutriente i. Se estivermos interessados numa dieta balanceada com uma ingestão mínima de calorias e além disso conhecermos a quantidade de calorias c_j de cada unidade/porção do alimento j, podemos

$$
\begin{array}{rl}
\min & c_1x_1 + \cdots + c_nx_n \\
\text{sujeito a} & a_{11}x_1 + \cdots + a_{1n}x_n \geq b_1 \\
& a_{21}x_1 + \cdots + a_{2n}x_n \geq b_2 \\
& \vdots \\
& a_{m1}x_1 + \cdots + a_{mn}x_n \geq b_m \\
& x \geq 0.
\end{array}
$$

[1]Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization.*

Seja $A \in \mathbb{R}^{m \times n}$ a matriz cujas entradas são a_{ij}. O mesmo problema pode ser escrito em notação matricial como

$$\begin{array}{rl} \min & c'x \\ \text{s.a} & Ax \geq b \\ & x \geq 0. \end{array}$$

Variante 1: Se estivermos interessados em obter a dieta balanceada mais barata e soubermos o preço r_j de cada unidade/porção do alimento j, então desejaremos minimizar a expressão $r'x$.

Variante 2: Se b representar os requisitos exatos de uma dieta "ideal", então x será uma dieta ideal se satisfizer $Ax = b,\ x \geq 0$. Sob estas restrições poderemos minimizar o conteúdo calórico ou o custo de uma dieta ideal.

Um problema de produção

Uma empresa produz n diferentes produtos usando m diferentes matérias-primas. Seja b_i a quantidade disponível da i-ésima matéria-prima, a_{ij} a quantidade de matéria-prima i necessária para a produção do produto j, e c_j o lucro obtido com a venda do produto j. Se a variável x_j representa a quantidade de produto j a ser produzida, então a empresa terá o máximo lucro resolvendo o problema

$$\begin{array}{rl} \max & c'x \\ \text{s.a} & Ax \leq b \\ & x \geq 0. \end{array}$$

O problema do plantão

Um hospital quer fazer a programação semanal dos plantões noturnos de seus enfermeiros. A cada dia da semana a demanda por enfermeiros de plantão é diferente, representada por um inteiro $d_j,\ j = 1, \ldots, 7$. Cada enfermeiro sempre trabalha 5 noites seguidas em plantão. O problema é encontrar o número mínimo de enfermeiros que o hospital precisa contratar.

Se tentássemos criar uma variável x_j representando o número de enfermeiros de plantão no dia j não seríamos capazes de escrever a restrição de que cada enfermeiro sempre trabalha 5 noites seguidas (experimente!). Ao invés disso, representamos por x_j o número de enfermeiros que *começa* a trabalhar no dia j; assim os enfermeiros que começarem a trabalhar no dia 5 trabalharão nos dias 5, 6, 7, 1 e 2. O problema pode então ser formulado como

$$
\begin{array}{lccccccccccccccccc}
\min & x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & + & x_6 & + & x_7 & & \\
\text{s.a} & x_1 & & & & & + & x_4 & + & x_5 & + & x_6 & + & x_7 & \geq & d_1 \\
& x_1 & + & x_2 & & & & & + & x_5 & + & x_6 & + & x_7 & \geq & d_2 \\
& x_1 & + & x_2 & + & x_3 & & & & & + & x_6 & + & x_7 & \geq & d_3 \\
& x_1 & + & x_2 & + & x_3 & + & x_4 & & & & & + & x_7 & \geq & d_4 \\
& x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & & & & & \geq & d_5 \\
& & + & x_2 & + & x_3 & + & x_4 & + & x_5 & + & x_6 & & & \geq & d_6 \\
& & & & & x_3 & + & x_4 & + & x_5 & + & x_6 & + & x_7 & \geq & d_7 \\
& x_j & \geq & 0, & & & & & x_j & \in & \mathbb{Z} & & & & &
\end{array}
$$

Exercício 1.1 *Verifique que a solução ótima da relaxação contínua deste problema (sem a restrição $x_j \in \mathbb{Z}$) pode ser obtida através da fórmula*

$$
x = \frac{1}{5}\begin{bmatrix}
3 & -2 & 3 & -2 & 3 & -2 & -2 \\
-2 & 3 & -2 & 3 & -2 & 3 & -2 \\
-2 & -2 & 3 & -2 & 3 & -2 & 3 \\
3 & -2 & -2 & 3 & -2 & 3 & -2 \\
-2 & 3 & -2 & -2 & 3 & -2 & 3 \\
3 & -2 & 3 & -2 & -2 & 3 & -2 \\
-2 & 3 & -2 & 3 & -2 & -2 & 3
\end{bmatrix} d.
$$

Este é um problema de programação linear *inteira*. Em algumas situações especiais um problema deste tipo pode ser resolvido como um problema de programação linear (sem a restrição $x_j \in \mathbb{Z}$).

Classificação de padrões

O problema de classificação de padrões corresponde a tentar identificar a classe de um objeto a partir da descrição de algumas de suas propriedades (seu padrão). Consideremos um exemplo bem simples: são dadas várias imagens de maçãs e laranjas, e para cada imagem um vetor $a \in \mathbb{R}^3$ tal que a_1 é a curvatura do objeto representado, a_2 é o comprimento da haste e a_3 é sua cor. O conjunto $\{a^i\}_{i \in S}$ contém padrões de maçãs, e $\{a^i\}_{i \notin S}$ contém padrões de laranjas. Um classificador linear para dintinguir as maçãs e laranjas dadas é um par $(x, y) \in \mathbb{R}^3 \times \mathbb{R}$ que satisfaz

$$
\begin{aligned}
(a^i)'x &\geq y, \quad i \in S \\
(a^i)'x &< y, \quad i \notin S.
\end{aligned}
$$

Dado um novo padrão $\bar{a}$ de uma imagem desconhecida, o mesmo será declarado pelo classificador uma maçã se satisfizer $\bar{a}'x \geq y$, uma laranja caso contrário.

Ordenação

Este exemplo ilustra a versatilidade de modelos de programação linear, mas não deve ser tomado como uma aplicação real: considere que se queira ordenar os números dados $a_1, \ldots, a_n$. O valor ótimo do problema abaixo

$$\begin{array}{ll} \min & a'x \\ \text{s.a} & \sum_{i=1}^{n} x_i = 1 \\ & x_i \in \{0,1\} \end{array}$$

é o menor dentre os valores $a_1, \ldots, a_n$.

Este é um exemplo de problema de programação linear *0-1* que pode ser reformulado como problema de programação linear trocando-se a restrição $x_i \in \{0,1\}$ por $0 \leq x_i \leq 1$. O método simplex (que veremos em breve) é capaz de encontrar soluções para o problema reformulado que satisfazem $x_i = 0$ ou $x_i = 1$.

Exercício 1.2 *Verifique que, trocando-se a restrição* $\sum_{i=1}^{n} x_i = 1$ *por* $\sum_{i=1}^{n} x_i = k$ *no problema original obtém-se como solução ótima a soma dos* k *menores valores dentre* $a_1, \ldots, a_n$.

Exercício 1.3 *Use o programa pl_solve para resolver o problema de planejamento de produção da DEC usando como entrada o arquivo abaixo:*

```
max: 60x1 + 40x2 + 30x3 + 30x4 + 15x5;

       x1 +   x2 +   x3 +   x4 +   x5 <= 7  ;
      4x1 +  2x2 +  2x3 +  2x4 +   x5 <= 8  ;
              x2        +   x4        <= 3  ;
       x1                             <= 1.8;
                     x3               <= 0.3;
       x1 +   x2 +   x3               <= 3.8;
                            x4 +   x5 <= 3.2;
              x2                      >= 0.5;
                            x4        >= 0.5;
                                   x5 >= 0.4;
```

Troque as restrições do problema original pelas restrições alternativas da modelagem e compare as soluções obtidas.

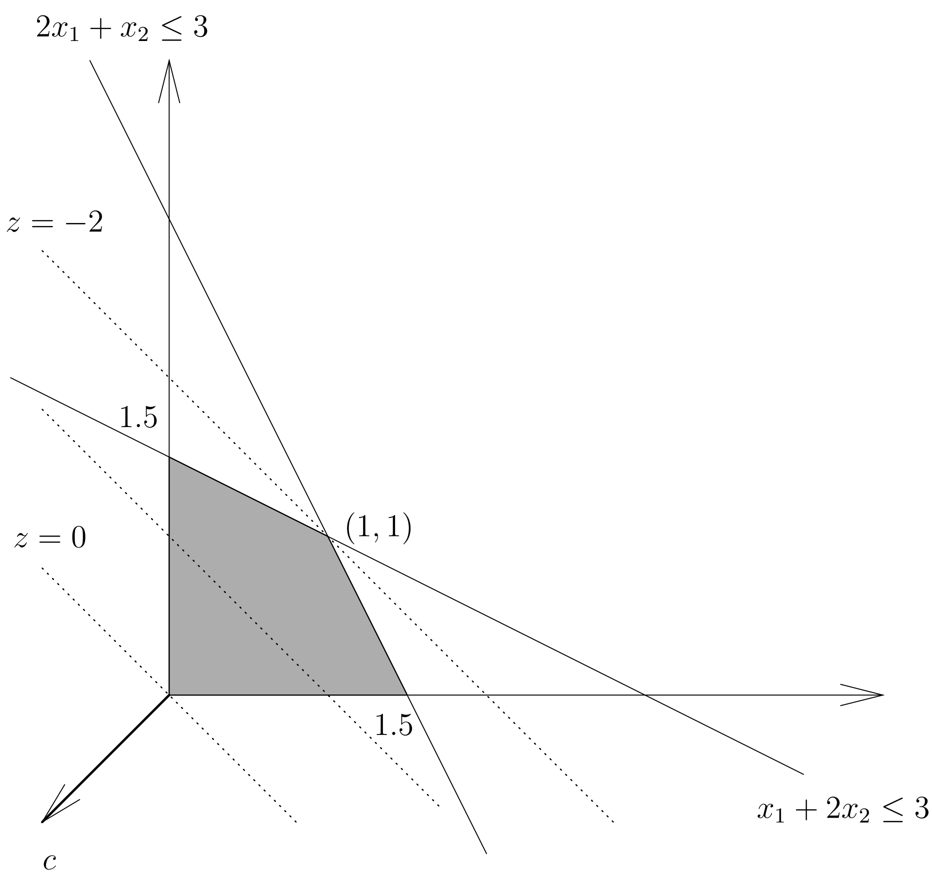
Resolva o problema do plantão usando o pl_solve (invente os valores de $d_1, \ldots, d_7$) e verifique que a solução obtida corresponde à fórmula do exercício 1.1.

1.2 Representação e solução gráficas

Considere o problema

$$\begin{array}{rrcrl} \min & -x_1 & - & x_2 & \\ \text{s.a} & x_1 & + & 2x_2 & \leq 3 \\ & 2x_1 & + & x_2 & \leq 3 \\ & & x_1, & x_2 & \geq 0 \end{array}$$

O conjunto viável está representado a seguir.



Para encontrar a solução ótima podemos considerar o conjunto dos pontos que têm um mesmo valor de função objetivo, digamos z. Este é uma linha descrita pela equação $-x_1 - x_2 = z$ e é perpendicular ao vetor $c = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$. Para cada valor de z obtemos uma linha paralela diferente:

aumentando z seguimos na direção apontada pelo vetor c, diminuindo z seguimos na direção $-c$. Para minimizar a função objetivo devemos procurar o ponto viável o mais distante na direção $-c$: este é o ponto $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ com valor ótimo $z = -2$.

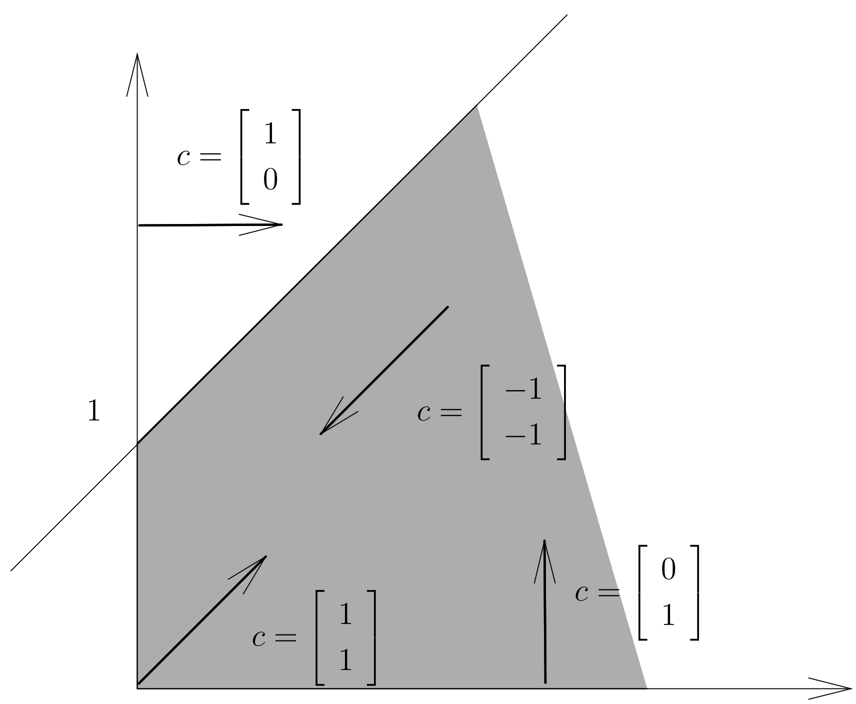
Uma prova algébrica de que esta é uma solução ótima é obtida ao notar-se que toda solução viável satisfaz $3x_1 + 3x_2 \leq 6$ (soma das duas primeiras restrições), ou equivalentemente, $-x_1 - x_2 \geq -2$. Assim x não só é viável como possui o menor valor possível dentre as soluções viáveis.

Exercício 1.4 *Obtenha graficamente a solução ótima do problema abaixo e verifique algebricamente que ela é de fato ótima.*

$$\begin{array}{rl} \min & -x_1 + x_2 - x_3 \\ s.a & 0 \leq x_1 \leq 1 \\ & 0 \leq x_2 \leq 1 \\ & 0 \leq x_3 \leq 1 \end{array}$$

Considere o exemplo a seguir:

$$\begin{array}{rcccl} \min & c_1x_1 & + & c_2x_2 & \\ \text{s.a} & -x_1 & + & x_2 & \leq 1 \\ & & x_1, & x_2 & \geq 0 \end{array}$$



1. Se $c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, a única solução ótima é $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

2. Se $c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, todo vetor da forma $x = \begin{bmatrix} 0 \\ x_2 \end{bmatrix}$ com $0 \leq x_2 \leq 1$ é solução ótima (infinitas soluções, conjunto de soluções limitado).

3. Se $c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, todo vetor da forma $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ com $x_1 \geq 0$ é solução ótima (infinitas soluções, conjunto de soluções ilimitado).

4. Se $c = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$, é possível obter uma seqüência de soluções viáveis com valores tendendo a $-\infty$. Dizemos que o problema é ilimitado e que o valor ótimo é $-\infty$ (mas nenhuma solução viável atinge este valor!)

5. Se o problema tivesse a restrição adicional $x_1 + x_2 \leq -2$, o conjunto viável seria vazio.

1.3 Variantes do problema de programação linear

Um problema de programação linear pode ser expresso da maneira mais geral como

$$
\begin{array}{rrcll}
\min / \max & c'x & & & \\
\text{s.a} & a_i'x & \geq & b_i, & i \in M_1, \\
 & a_i'x & \leq & b_i, & i \in M_2, \\
 & a_i'x & = & b_i, & i \in M_3, \\
 & x_j & \geq & 0, & j \in N_1, \\
 & x_j & \leq & 0, & j \in N_2.
\end{array}
$$

$x_1, \ldots, x_n$ são chamadas *variáveis de decisão* e um vetor x que satisfaz todas as restrições é chamado de *solução viável* ou *ponto viável.* O conjunto de soluções viáveis é chamado de *conjunto viável* ou *região viável.* Para $j \notin N_1 \cup N_2$ dizemos que x_j é uma variável *livre de sinal.* A função $f(x) = c'x$ é chamada *função objetivo* e uma solução viável x^* que minimiza a função objetivo (isto é, que satisfaz $c'x^* \leq c'x$ para todo x viável) é chamada de *solução ótima* do problema de minimização, com o correspondente *valor ótimo* igual a $c'x^*$ (define-se analogamente solução ótima do problema de maximização). Se, por outro lado, for possível encontrar soluções viáveis com valores de função objetivo tão baixos quanto se queira, diz-se que o valor ótimo é $-\infty$ e que o problema é ilimitado.

Um tal grau de generalidade na formulação é inconveniente e desnecessário do ponto de vista da teoria, que pode ser desenvolvida para moldes mais simples de problemas lineares. Discutimos a seguir algumas reformulações que não tiram expressividade do modelo de programação linear.

$$\begin{array}{c} x^* \text{ é sol. ótima de } \max c'x \\ \Updownarrow \\ c'x^* \geq c'x \ \forall x \text{ viável} \\ \Updownarrow \\ -c'x^* \leq -c'x \ \forall x \text{ viável} \\ \Updownarrow \\ x^* \text{ é sol. ótima de } \min -c'x. \end{array}$$

Inicialmente note que não é necessário estudar problemas de maximização e de minimização separadamente, pois são equivalentes $\max c'x$ e $\min -c'x$.

Toda restrição de igualdade do tipo $a_i'x = b_i$ é equivalente às duas desigualdades $a_i'x \leq b_i$ e $a_i'x \geq b_i$; além disso qualquer desigualdade do tipo $a_i'x \leq b_i$ pode ser re-escrita como $(-a_i)'x \geq -b_i$. Note ainda que as restrições de sinal $x_j \geq 0$ e $x_j \leq 0$ são casos particulares das restrições do tipo $a_i'x \geq b_i$, onde $a_i = e_j$ e $b_i = 0$.

Através destas equivalências concluimos que qualquer problema de programação linear pode ser escrito na forma compacta que designamos de *forma geral de programação linear*:

$$\begin{array}{rl} \min & c'x \\ \text{s.a} & Ax \geq b. \end{array}$$

Outra forma muito importante no desenvolvimento de algoritmos de programação linear é a *forma canônica (ou padrão) de programação linear*:

$$\begin{array}{rl} \min & c'x \\ \text{s.a} & Ax = b \\ & x \geq 0. \end{array}$$

Vimos que a forma canônica é um caso particular da forma geral, ou seja, podemos re-escrever qualquer problema em forma canônica na forma geral. O inverso também é verdade: pode-se transformar qualquer problema de programação linear para a forma canônica. Esta transformação depende de duas operações básicas:

Eliminação de variáveis livres de sinal Dada uma variável x_j livre de sinal, substitui-se todas as ocorrências desta por $x_j^+ - x_j^-$, onde x_j^+, x_j^- são novas variáveis não-negativas, isto é, sujeitas às condições $x_j^+ \geq 0$ e $x_j^- \geq 0$. A idéia é que todo número real pode ser escrito como a diferença de dois números não-negativos.

Eliminação de desigualdades Dada uma desigualdade da forma $a_i'x \leq b_i$, introduz-se uma nova variável r_i e as restrições (compatíveis com a forma canônica)

$$\begin{array}{c} a_i'x + r_i = b_i \\ r_i \geq 0. \end{array}$$

A variável r_i é chamada de *residual*.

Naturalmente uma variável não-positiva $x_j \leq 0$ é substituida por $-x_j$, e uma desigualdade da forma $a_i'x \geq b_i$ é substituida por $a_i'x - r_i = b_i$ e $r_i \geq 0$.

Como exemplo, o problema

$$\begin{array}{rrcrcrcrcr}
\min & 2x_1 & - & x_2 & + & 4x_3 & & & & \\
\text{s.a} & x_1 & + & x_2 & & & + & x_4 & \leq & 2 \\
 & & & 3x_2 & - & x_3 & & & = & 5 \\
 & & & & & x_3 & + & x_4 & \geq & 3 \\
 & x_1 & & & & & & & \geq & 0 \\
 & & & & & x_3 & & & \leq & 0
\end{array}$$

pode ser re-escrito na forma geral como

$$\begin{array}{rrcrcrcrcr}
\min & 2x_1 & - & x_2 & + & 4x_3 & & & & \\
\text{s.a} & -x_1 & - & x_2 & & & - & x_4 & \geq & -2 \\
 & & & 3x_2 & - & x_3 & & & \geq & 5 \\
 & & & -3x_2 & + & x_3 & & & \geq & -5 \\
 & & & & & x_3 & + & x_4 & \geq & 3 \\
 & x_1 & & & & & & & \geq & 0 \\
 & & & & & -x_3 & & & \geq & 0
\end{array}$$

e na forma canônica como

$$\begin{array}{rrcrcrcrcrcrcrcr}
\min & 2x_1 & - & x_2^+ & + & x_2^- & - & 4x_3 & & & & & & & & \\
\text{s.a} & x_1 & + & x_2^+ & - & x_2^- & & & + & x_4 & + & r_1 & & & = & 2 \\
 & & & 3x_2^+ & - & 3x_2^- & + & x_3 & & & & & & & = & 5 \\
 & & & & & & & -x_3 & + & x_4 & & & - & r_2 & = & 3 \\
 & x_1, & & x_2^+, & & x_2^-, & & x_3, & & x_4, & & r_1, & & r_2 & \geq & 0.
\end{array}$$

Exercício 1.5 *Passe o problema de planejamento da DEC para a forma canônica.*

1.4 Pré-requisitos e notação

Conjuntos

Utiliza-se as seguintes notações: $|S|$ denota a cardinalidade do conjunto S, $S \backslash T = \{s \in S \mid s \notin T\}$, $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ e $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$.

Vetores e matrizes

$A \in \mathbb{R}^{m\times n}$ denota uma matriz

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

$x \in \mathbb{R}^n$ denota o vetor coluna $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$; alguns vetores especiais utilizados são $0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ e e^i o i-ésimo vetor da base canônica.

A' denota a *transposta* de A, que satisfaz $A'_{ij} = A_{ji}$, $\forall i, j$. Se $x, y \in \mathbb{R}^n$, então $x'y = \sum_{i=1}^n x_i y_i$. Lembre-se que $x \perp y$ (x é ortogonal a y) $\iff$ $x'y = 0$. A *norma Euclideana* de $x \in \mathbb{R}^n$ é denotada por $\|x\| = \sqrt{x'x}$.

Exercício 1.6 *Demonstre a desigualdade de Cauchy-Schwarz:*

$$|x'y| \leq \|x\|\|y\|,$$

e verifique que a igualdade vale se, e somente se, x é múltiplo de y ou vice-versa. Dica: se $x \neq 0$, considere a distância de y à projeção de y sobre x.

Dada uma matriz A denota-se por A^j sua j-ésima coluna, e A_i sua i-ésima linha. O *produto* de duas matrizes $A \in \mathbb{R}^{m\times n}$ e $B \in \mathbb{R}^n \times k$ é uma matriz em $\mathbb{R}^{m\times k}$ cujas componentes são dadas por $[AB]_{ij} = \sum_{l=1}^n a_{il} b_{lj} = A_i B^j$, ou ainda,

$$AB = \left[\begin{array}{c|c|c} AB^1 & \cdots & AB^k \end{array}\right] = \left[\begin{array}{c} A_1 B \\ \hline \vdots \\ \hline A_m B \end{array}\right].$$

O produto de matrizes satisfaz $(AB)C = A(BC)$ e $(AB)' = B'A'$, mas não satisfaz $AB = BA$ em geral (produza um contra-exemplo!) A matriz *identidade* é denotada por I e satisfaz $IA = A$ e $BI = B$ para quaisquer A e B de dimensões compatíveis.

Dados $A \in \mathbb{R}^{m\times n}$ e $x \in \mathbb{R}^n$, valem as identidades $Ae^i = A^i$, $x = \sum_{i=1}^n x_i e^i$ e

$$Ax = \sum_{i=1}^{n} A^i x_i = \begin{bmatrix} A_1 x \\ \vdots \\ A_m x \end{bmatrix}.$$

Dado $x \in \mathbb{R}^n$ utiliza-se as notações $x \geq 0 \iff x_i \geq 0\ \forall i$ e $x > 0 \iff x_i > 0\ \forall i$; analogamente define-se as desigualdades $A \geq 0$ e $A > 0$. Estas definições estendem-se naturalmente: $x \geq y \iff x - y \geq 0$ e $x > y \iff x - y > 0$.

Inversão de matrizes

Uma matriz quadrada A é dita *não-singular* ou *inversível* se existe uma matriz B tal que $AB = BA = I$; esta matriz é única e denotada por A^{-1}. Se $A, B \in \mathbb{R}^{n\times n}$ são inversíveis então AB é inversível e $(AB)^{-1} = B^{-1}A^{-1}$.

Uma coleção de vetores $x^1, \ldots, x^k \in \mathbb{R}^n$ é dita *linearmente dependente* se existem $a_1, \ldots, a_k$ reais tais que $a \neq 0$ e $\sum_{i=1}^k a_i x^i = 0$; caso contrário diz-se que os vetores são *linearmente independentes*. Vale o seguinte resultado

Teorema 1.2 *Seja A uma matriz quadrada. As afirmações seguintes são equivalentes:*

1. *A é inversível;*
2. *A' é inversível;* $A'(A^{-1})' = (A^{-1}A)' = I' = I$, $(A^{-1})'A' = (AA^{-1})' = I' = I$.
3. $\det(A) \neq 0$*;* $\det(A^{-1})\det(A) = \det(I) = 1$.
4. *$A_1, \ldots, A_n$ são linearmente independentes;*
5. *$A^1, \ldots, A^n$ são linearmente independentes;*
6. *Existe b tal que o sistema $Ax = b$ tem uma única solução.* $Ax = 0$ possui solução única.
7. *Para todo b o sistema $Ax = b$ tem uma única solução;*

Prova.

$(3 \Longrightarrow 4)$ Suponha que $\sum_{i=1}^{n} \alpha_i A_i = 0$, e considere

$$B = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Então $[BA]_1 = \sum_{i=1}^{n} \alpha_i A_i = 0 \Longrightarrow 0 = \det(BA) = \det(B)\det(A) \Longrightarrow \det(B) = \alpha_1 = 0$. Repita o argumento para $\alpha_2, \ldots, \alpha_n$.

$(4 \Longrightarrow 5)$ Suponha que $A\alpha = \sum_{i=1}^{n} \alpha_i A^i = 0$. Então $A_i \alpha = 0,\ \forall i$, logo $\alpha = \sum_{i=1}^{n} \mathrm{proj}_{A_i}(\alpha) = \frac{A_i \alpha}{\|A_i\|^2} A_i = 0$.

$(6 \Longrightarrow 7)$ Suponha que $Ax = b$ possui uma única solução z e que $Ax = b'$ possui duas soluções x e y. Então $A(z + x - y) = Az + A(x - y) = Az + Ax - Ay = b - b' - b' = b$, logo $z + x - y$ é solução de $Ax = b$ e portanto $x = y$.

$(7 \Longrightarrow 1)$ Defina $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ por $T(b) = x \iff Ax = b$. Então $AT(b) = b,\ \forall b \in \mathbb{R}^n$ e $TA(x) = x,\ \forall x \in \mathbb{R}^n$, ou seja, $AT \equiv TA \equiv I$ e a matriz que representa T é a inversa de A.

Subespaços e bases

$\emptyset \neq S \subset \mathbb{R}^n$ é um *subespaço* de $\mathbb{R}^n$ se $ax + by \in S$ para quaisquer $x, y \in S$ e $a, b \in \mathbb{R}$. O *subespaço gerado* por $x^1, \ldots, x^k$ é o conjunto de vetores y da forma $y = \sum_{i=1}^{k} a_i x^i$, onde $a \in \mathbb{R}^k$; y é dito uma *combinação linear* de $x^1, \ldots, x^k$.

Se B e C geram S, e $n = |C| > |B| = m$, então $C = BA$ onde $A \in \mathbb{R}^{m\times n}$. Seja $x \neq 0 : Ax = 0$. Então $Cx = BAx = B0 = 0$, ou seja, C não é LI.

Dado um subespaço $\emptyset \neq S \subset \mathbb{R}^n$, uma *base* de S é uma coleção de vetores linearmente independentes que geram S. Toda base de S tem o mesmo número de vetores e este número é chamado de *dimensão* de S; em particular $\dim(\mathbb{R}^n) = n$ e todo subespaço de $\mathbb{R}^n$ tem dimensão menor ou igual a n. Por definição, $\dim(\{0\}) = 0$.

Se S é um subespaço *próprio* de $\mathbb{R}^n$ (isto é, diferente de $\mathbb{R}^n$), então existe um vetor $a \neq 0$ tal que $a'x = 0$ para todo $x \in S$. Em geral, se $\dim(S) = m < n$, então existem $n - m$ vetores linearmente independentes ortogonais a S.

Teorema 1.3 *Suponha que o subespaço S gerado pelos vetores $x^1, \ldots, x^k$ tem dimensão m. Então:*

1. *Existe uma base de S contendo m vetores dentre $x^1, \ldots, x^k$;*

2. *Se $l \leq m$ e $x^1, \ldots, x^l$ são linearmente independentes, pode-se formar uma base de S começando com $x^1, \ldots, x^l$ e escolhendo $m - l$ vetores dentre $x^{l+1}, \ldots, x^k$*

Exercício 1.7 *Demonstre o teorema acima. Note que a parte 1 é caso particular da parte 2 (com $l = 0$).*

Seja $A \in \mathbb{R}^{m \times n}$. O *espaço-coluna* gerado por A é o subespaço (de $\mathbb{R}^m$) gerado pelas colunas de A; analogamente definimos o *espaço-linha* de A. As dimensões dos espaço-linha e espaço-coluna de A coincidem e este número é denominado *posto* ou *característica* de A. A possui *posto completo* ou *característica plena* se $\text{posto}(A) = \min\{m, n\}$. O conjunto $\{x \in \mathbb{R}^n \mid Ax = 0\}$ é chamado de *espaço-nulo* de A.

Exercício 1.8 *Prove que*

1. $\dim(\textit{nulo}(A)) = n - \dim(\textit{linhas}(A))$.

2. $\dim(\textit{colunas}(A)) + \dim(\textit{nulo}(A)) = n$.

Observe que 1 e 2 implicam $\dim(\textit{col}(A)) = \dim(\textit{lin}(A))$.

Subespaços afins

Dado um subespaço linear $S_0 \subset \mathbb{R}^n$ e um vetor $x_0 \in \mathbb{R}^n$, o conjunto $S = S_0 + x_0 = \{x + x_0 \mid x \in S_0\}$ corresponde a uma translação do subespaço S_0 e é denominado *subespaço afim*. S tem a mesma dimensão de S_0, por definição.

Como um primeiro exemplo, considere $k + 1$ vetores em $\mathbb{R}^n$, e o conjunto $S = \{x^0 + \lambda_1 x^1 + \cdots + \lambda_k x^k \mid \lambda \in \mathbb{R}^k\}$. S é um subespaço afim obtido a partir da translação do subespaço S_0 gerado pelos vetores $x^1, \ldots, x^k$. Se estes últimos vetores forem linearmente independentes, então tanto S_0 quanto S terão dimensão k.

Outro exemplo é dado pelo conjunto $S = \{x \in \mathbb{R}^n \mid Ax = b\}$ onde $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ e supomos $S \neq \emptyset$. Seja $x^0 \in \mathbb{R}^n$ tal que $Ax^0 = b$; então $x \in S \iff Ax = Ax^0 = b$, ou seja, $x \in S \iff A(x - x^0) = 0 \iff x - x^0 \in S_0 = \{y \mid Ay = 0\}$. Ou seja, S é a translação do espaço-nulo de A pelo vetor x^0. Cada restrição $a_i' x = b_i$ define um hiperplano de $\mathbb{R}^n$ (um subespaço de dimensão $n - 1$), e em princípio diminui em 1 a dimensão do conjunto S: se A possui m linhas linearmente independentes, então $\dim(S) = n - m$.